

# Raw Sequence Listing Error Summary

## ERROR DETECTED   SUGGESTED   CORRECTION

SERIAL NUMBER: 09/144,838

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleiks      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Amino Acids      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.  
All text must be visible on page.
- 4        Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) features section that some may be missing.
- 7        Wrong Designation      Sequence(s)        contain amino acid or nucleic acid designators which are not standard  
representations as per the Sequence Rules (Please refer to paragraph 1.822)
- 8        Skipped Sequences (OLD RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(I) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences (NEW RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10        Use of N's or Xaa's (NEW RULES)      Use of N's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
- 11        Use of <213>Organism (NEW RULES)      Sequence(s)        are missing this mandatory field or its response.
- 12        Use of <220>Feature (NEW RULES)      Sequence(s)        are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)  
(Sec. 1.823 of new Sequence Rules)
- 13        Wrong Format      File submitted was in the alphabetical heading format of the Old Sequence Rules. This is invalid since the  
"Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures"  
Federal Register Notice, Vol. 63, No. 104, June 1, 1998, p. 29620  
applies to applications filed on or after July 1, 1998.

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/144,838

DATE: 12/01/1998  
TIME: 12:57:03

Input Set: I144838.RAW

This Raw Listing contains the General Information  
Section and up to first 5 pages.

1 <110> APPLICANT: Siani, Michael A.  
2 Wilken, Jill  
3 Simon, Reyna  
4 Kent, Stephen B.H.  
5 <120> TITLE OF INVENTION: Modular Protein Libraries and Methods of Preparation  
6 <130> FILE REFERENCE: GRFN20/O1US  
7 <140> CURRENT APPLICATION NUMBER: US/09/144,838  
8 <141> CURRENT FILING DATE: 1998-08-31  
9 <150> EARLIER APPLICATION NUMBER: US 60/057,620  
10 <151> EARLIER FILING DATE: 1997-09-04  
11 <160> NUMBER OF SEQ ID NOS: 54  
12 <170> SOFTWARE: PatentIn Ver. 2.0  
13 <210> SEQ ID NO 1  
14 <211> LENGTH: 81  
15 <212> TYPE: PRT  
16 <213> ORGANISM: Artificial Sequence  
17 <220> FEATURE:  
18 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
19 <400> SEQUENCE: 1

W--> 20 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Leu Ile Val Met Xaa  
21 1 5 10 15  
W--> 22 Xaa Xaa Xaa Xaa Xaa Leu Ile Val Met Phe Tyr Xaa Xaa Arg Lys Ser  
23 20 25 30  
W--> 24 Glu Gln Xaa Leu Ile Val Met Xaa Xaa Leu Ile Val Met Xaa Xaa Xaa  
25 35 40 45  
W--> 26 Xaa Xaa Ser Ala Gly Xaa Xaa Cys Xaa Xaa Xaa Glu Gln Leu Ile Val  
27 50 55 60  
W--> 28 Met Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Leu Asp  
29 65 70 75 80  
30 Asn

31 <210> SEQ ID NO 2  
32 <211> LENGTH: 82  
33 <212> TYPE: PRT  
34 <213> ORGANISM: Artificial Sequence  
35 <220> FEATURE:  
36 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
37 <400> SEQUENCE: 2

W--> 38 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Cys Leu Ile Val Met Phe  
39 1 5 10 15  
W--> 40 Tyr Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Ile Val Met Xaa Xaa Xaa Xaa  
41 20 25 30  
W--> 42 Leu Ile Val Met Phe Xaa Xaa Tyr Xaa Xaa Xaa Gly Ser Thr Asn Gly  
43 35 40 45  
W--> 44 Ser Thr Asn Xaa Xaa Cys Xaa Xaa Xaa Xaa Ser Ala Gly Leu Ile Val

New format

Does Not Comply  
Corrected Diskette Needed

See  
Item 10  
on Enon  
summary  
sheet

See  
Item 10

PAGE: 2

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION US/09/144,838**

DATE: 12/01/1998

TIME: 12:57:03

Input Set: I144838.RAW

```

45             50             55             60
W--> 46      Met Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa Arg Lys Thr Met Phe Xaa
47             65             70             75             80
W--> 48      Xaa Cys
49      <210> SEQ ID NO 3
50      <211> LENGTH: 71
51      <212> TYPE: PRT
52      <213> ORGANISM: Artificial Sequence
53      <220> FEATURE:
54      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
55      <400> SEQUENCE: 3
56      Ala Gly Ser Leu Val Ser Tyr Thr Pro Asn Ser Cys Cys Tyr Gly Phe
57             1             5             10             15
58      Gln Gln His Pro Pro Pro Val Gln Ile Leu Lys Glu Trp Tyr Pro Thr
59             20             25             30
60      Ser Pro Ala Cys Pro Lys Pro Gly Val Ile Leu Leu Thr Lys Arg Gly
61             35             40             45
62      Arg Gln Ile Cys Ala Asp Pro Ser Lys Asn Trp Val Arg Gln Leu Met
63             50             55             60
64      Gln Arg Leu Pro Ala Ile Ala
65             65             70
66      <210> SEQ ID NO 4
67      <211> LENGTH: 74
68      <212> TYPE: PRT
69      <213> ORGANISM: Artificial Sequence
70      <220> FEATURE:
71      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
72      <400> SEQUENCE: 4
73      Gly Asp Thr Leu Gly Ala Ser Trp His Arg Pro Asp Lys Cys Cys Leu
74             1             5             10             15
75      Gly Tyr Gln Lys Arg Pro Leu Pro Gln Val Leu Leu Ser Ser Trp Tyr
76             20             25             30
77      Pro Thr Ser Gln Leu Cys Ser Lys Pro Gly Val Ile Phe Leu Thr Lys
78             35             40             45
79      Arg Gly Arg Gln Val Cys Ala Asp Lys Ser Lys Asp Trp Val Lys Lys
80             50             55             60
81      Leu Met Gln Gln Leu Pro Val Thr Ala Arg
82             65             70
83      <210> SEQ ID NO 5
84      <211> LENGTH: 72
85      <212> TYPE: PRT
86      <213> ORGANISM: Artificial Sequence
87      <220> FEATURE:
88      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
89      <400> SEQUENCE: 5
90      Ala Gly Ser Leu Val Ser Tyr Thr Pro Asn Ser Cys Cys Tyr Gly Phe
91             1             5             10             15
92      Gln Gln His Pro Pro Pro Val Gln Ile Leu Lys Glu Trp Tyr Pro Thr
93             20             25             30
94      Ser Pro Ala Cys Ser Lys Pro Gly Val Ile Phe Leu Thr Lys Arg Gly

```

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/144,838

DATE: 12/01/1998  
TIME: 12:57:03

Input Set: I144838.RAW

```

95          35          40          45
96      Arg Gln Val Cys Ala Asp Lys Ser Lys Asp Trp Val Lys Lys Leu Met
97          50          55          60
98      Gln Gln Leu Pro Val Thr Ala Arg
99          65          70
100 <210> SEQ ID NO 6
101 <211> LENGTH: 73
102 <212> TYPE: PRT
103 <213> ORGANISM: Artificial Sequence
104 <220> FEATURE:
105 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
106 <400> SEQUENCE: 6
107      Gly Asp Thr Leu Gly Ala Ser Trp His Arg Pro Asp Lys Cys Cys Leu
108          1          5          10          15
109      Gly Tyr Gln Lys Arg Pro Leu Pro Gln Val Leu Leu Ser Ser Trp Tyr
110          20          25          30
111      Pro Thr Ser Gln Leu Cys Pro Lys Pro Gly Val Ile Leu Leu Thr Lys
112          35          40          45
113      Arg Gly Arg Gln Ile Cys Ala Asp Pro Ser Lys Asn Trp Val Arg Gln
114          50          55          60
115      Leu Met Gln Arg Leu Pro Ala Ile Ala
116          65          70
117 <210> SEQ ID NO 7
118 <211> LENGTH: 93
119 <212> TYPE: PRT
120 <213> ORGANISM: Artificial Sequence
121 <220> FEATURE:
122 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
123 <400> SEQUENCE: 7
124      Met Asn Ala Lys Val Val Val Val Leu Val Leu Val Leu Thr Ala Leu
125          1          5          10          15
126      Cys Leu Ser Asp Gly Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys
127          20          25          30
128      Arg Phe Phe Glu Ser His Val Ala Arg Ala Asn Val Lys His Leu Lys
129          35          40          45
130      Ile Leu Asn Thr Pro Asn Cys Ala Leu Gln Ile Val Ala Arg Leu Lys
131          50          55          60
132      Asn Asn Asn Arg Gln Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln
133          65          70          75          80
134      Glu Tyr Leu Glu Lys Ala Leu Asn Lys Arg Phe Lys Met
135          85          90
136 <210> SEQ ID NO 8
137 <211> LENGTH: 67
138 <212> TYPE: PRT
139 <213> ORGANISM: Artificial Sequence
140 <220> FEATURE:
141 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
142 <400> SEQUENCE: 8
143      Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys Arg Phe Phe Glu Ser
144          1          5          10          15

```

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/144,838

DATE: 12/01/1998  
TIME: 12:57:03

Input Set: I144838.RAW

```

145      His Val Ala Arg Ala Asn Val Lys His Leu Lys Ile Leu Asn Thr Pro
146              20              25              30
147      Asn Cys Ala Leu Gln Ile Val Ala Arg Leu Lys Asn Asn Asn Arg Gln
148              35              40              45
149      Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln Glu Tyr Leu Glu Lys
150              50              55              60
151      Ala Leu Asn
152      65
153      <210> SEQ ID NO 9
154      <211> LENGTH: 91
155      <212> TYPE: PRT
156      <213> ORGANISM: Artificial Sequence
157      <220> FEATURE:
158      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
159      <400> SEQUENCE: 9
160      Met Lys Val Ser Ala Ala Arg Leu Ala Val Ile Leu Ile Ala Thr Ala
161      1              5              10              15
162      Leu Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro
163              20              25              30
164      Cys Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys
165              35              40              45
166      Glu Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val Phe
167              50              55              60
168      Val Thr Arg Lys Asn Arg Gln Val Cys Ala Asn Pro Glu Lys Lys Trp
169              65              70              75              80
170      Val Arg Glu Tyr Ile Asn Ser Leu Glu Met Ser
171              85              90
172      <210> SEQ ID NO 10
173      <211> LENGTH: 68
174      <212> TYPE: PRT
175      <213> ORGANISM: Artificial Sequence
176      <220> FEATURE:
177      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic
178      <400> SEQUENCE: 10
179      Ser Pro Tyr Ser Ser Asp Thr Thr Pro Cys Cys Phe Ala Tyr Ile Ala
180      1              5              10              15
181      Arg Pro Leu Pro Arg Ala His Ile Lys Glu Tyr Phe Tyr Thr Ser Gly
182              20              25              30
183      Lys Cys Ser Asn Pro Ala Val Val Phe Val Thr Arg Lys Asn Arg Gln
184              35              40              45
185      Val Cys Ala Asn Pro Glu Lys Lys Trp Val Arg Glu Tyr Ile Asn Ser
186              50              55              60
187      Leu Glu Met Ser
188      65
189      <210> SEQ ID NO 11
190      <211> LENGTH: 33
191      <212> TYPE: PRT
192      <213> ORGANISM: Artificial Sequence
193      <220> FEATURE:
194      <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic

```

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/144,838

DATE: 12/01/1998

TIME: 12:57:03

Input Set: I144838.RAW

195 <400> SEQUENCE: 11  
196 Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys Arg Phe Phe Glu Ser  
197 1 5 10 15  
198 His Val Ala Arg Ala Asn Val Lys His Leu Lys Ile Leu Asn Thr Pro  
199 20 25 30  
200 Asn  
201 <210> SEQ ID NO 12  
202 <211> LENGTH: 32  
203 <212> TYPE: PRT  
204 <213> ORGANISM: Artificial Sequence  
205 <220> FEATURE:  
206 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
207 <400> SEQUENCE: 12  
208 Lys Pro Val Ser Leu Ser Tyr Arg Cys Cys Arg Phe Phe Glu Ser His  
209 1 5 10 15  
210 Val Ala Arg Ala Asn Val Lys His Leu Lys Ile Leu Asn Thr Pro Asn  
211 20 25 30  
212 <210> SEQ ID NO 13  
213 <211> LENGTH: 33  
214 <212> TYPE: PRT  
215 <213> ORGANISM: Artificial Sequence  
216 <220> FEATURE:  
217 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
218 <400> SEQUENCE: 13  
219 Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys Phe Ala Tyr Ile Ala  
220 1 5 10 15  
221 Arg Pro Leu Pro Arg Ala His Ile Lys Glu Tyr Phe Tyr Thr Ser Gly  
222 20 25 30  
223 Lys  
224 <210> SEQ ID NO 14  
225 <211> LENGTH: 32  
226 <212> TYPE: PRT  
227 <213> ORGANISM: Artificial Sequence  
228 <220> FEATURE:  
229 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
230 <400> SEQUENCE: 14  
231 Lys Pro Val Ser Leu Ser Tyr Arg Cys Cys Phe Ala Tyr Ile Ala Arg  
232 1 5 10 15  
233 Pro Leu Pro Arg Ala His Ile Lys Glu Tyr Phe Tyr Thr Ser Gly Lys  
234 20 25 30  
235 <210> SEQ ID NO 15  
236 <211> LENGTH: 33  
237 <212> TYPE: PRT  
238 <213> ORGANISM: Artificial Sequence  
239 <220> FEATURE:  
240 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
241 <400> SEQUENCE: 15  
242 Ser Pro Tyr Ser Ser Asp Thr Thr Pro Cys Cys Phe Ala Tyr Ile Ala  
243 1 5 10 15  
244 Arg Pro Leu Pro Arg Ala His Ile Lys Glu Tyr Phe Tyr Thr Ser Gly

VERIFICATION SUMMARY  
PATENT APPLICATION US/09/144,838DATE: 12/01/1998  
TIME: 12:57:03

Input Set: I144838.RAW

Line	Error/Warning	Original Text
20	W "N" or "Xaa" used: Feature required	Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa C
22	W "N" or "Xaa" used: Feature required	Xaa Xaa Xaa Xaa Xaa Leu Ile Val Met Phe T
24	W "N" or "Xaa" used: Feature required	Glu Gln Xaa Leu Ile Val Met Xaa Xaa Leu I
26	W "N" or "Xaa" used: Feature required	Xaa Xaa Ser Ala Gly Xaa Xaa Cys Xaa Xaa X
28	W "N" or "Xaa" used: Feature required	Met Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa X
38	W "N" or "Xaa" used: Feature required	Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys C
40	W "N" or "Xaa" used: Feature required	Tyr Thr Xaa Xaa Xaa Xaa Xaa Xaa Leu Ile V
42	W "N" or "Xaa" used: Feature required	Leu Ile Val Met Phe Xaa Xaa Tyr Xaa Xaa X
44	W "N" or "Xaa" used: Feature required	Ser Thr Asn Xaa Xaa Cys Xaa Xaa Xaa Xaa S
46	W "N" or "Xaa" used: Feature required	Met Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa A
48	W "N" or "Xaa" used: Feature required	Xaa Cys